804 Tastes: Evidence on Preferences, Randomness, and Value from Double-Blind Wine Tastings*

Jeffrey C. Bodington [a]

Abstract

Results for a total of 804 double-blind tastes by experienced tasters during nine tasting events are reported. *T*-test results reject the hypothesis that flight-position bias affects results. The distribution of ranks for a wine is a mixture distribution, and tests concerning the variance of that mixture distribution do not isolate the variance due to the randomness mixture component alone. *T*-statistics for the mean ranks of high- and low-ranking wines are over several standard deviations from a random expectation. *T*-tests show that the statistical significance of the difference between wine ranks is positively related to the difference in their mean ranks. At a 95% level of significance, the difference in ranks between the first- and second-place wines appears to be significant in 33% of tastings. At 95%, the difference in ranks between the first- and last-place wines appears to be significant in 100% of tastings. Monte Carlo simulation shows that much of those differences could be illusory and due to ranking procedures that lead to Type I errors. While the mean correlation coefficient between price per bottle and mean preference is a weakly positive 0.23, this may not indicate an inefficient market. (JEL Classifications: A10, C00, C12, D12)

I. Introduction

Wine tastings can inspire lively discussion, commercial insight, controversy, and occasional offense. Some have found evidence that tasting results are significantly random (Weil, 2001, 2005). Others have found that tasting results are influenced by expertise, preferences, and many other factors, including the sex of a taster

(Almenberg and Dreber, 2009; Amerine and Roessler, 1976; Castriota, Curzi, and Delmastro, 2012; Lecocq and Visser, 2006; Pecotich and Ward, 2010). Even wine competition judges' assessments to some extent appear random (Ashenfelter, 2006; Cliff and King, 1997; Hodgson, 2008; Quandt, 2012). On that foundation, it is not surprising that several analysts have shown that a wine's price is only a rough guide to how it compares to other wines in a flight (Almenberg and Dreber, 2009; Goldstein et al., 2008; Pecotich and Ward, 2010).

   Data from double-blind wine tastings presented herein deepens our understanding of tasting results and in some respects contradicts the results noted above. In certain respects, *t*-tests show that tasting results are not random at a 95% level of significance. The subject tasting data come from a group founded by the late wine writer Stephen R. Pitcher and collector David S. Rosen during the late 1980s.[1] The group remains active and has tasted, ranked, and discussed over 1,200 wines during the past 25 years. The implications of nine recent tastings are the subject of this article.[2] The group's tasting procedure is described in section II, analysis of the tasting results appears in section III, and conclusions follow in section IV.

II. Double-Blind Tasting Protocol

Each tasting involved six wines and nine to 19 participants, including wine writers, makers, collectors, distributors, and enthusiasts. In advance of each tasting, one participant placed each wine in a plain brown paper bag and removed both the cork and identifying neck foil, if any. A second participant, who did not see the wines as they were bagged, then mixed up the bagged bottles and labeled each with one of the letters A through F. Tastes were poured for each member, from left to right, in each member's six glasses. Along with the pour, each taster received a score sheet on which to make notes about and score each wine.

   Next, a sponsoring member provided the group with an overview of the wines. This overview included the wines' variety and the ranges of years, appellations, and prices. In general, each tasting involved one variety from either one or contrasting appellations and a narrow range of years. Prices usually covered a broad range, however, there were no low-cost "ringers" in the subject data. Specific wine makers, vineyards, and prices were never mentioned in opening remarks. That information was disclosed only after all the scores had been tabulated. After the opening remarks, the participants then tasted and scored their wines. Discussion during this tasting period was minimal. Tasters ultimately ranked each wine from most-favored

---

[1] The author was an early member of this group and remains active.
[2] Data on each tasting include each taster's rank and the per-bottle price of each wine.

first place to least-favored sixth place and handed their score sheets to one member, who recorded the results.

After all tasters' results were recorded, the recording member calculated each wine's total of ranks, mean of ranks, and the number of rankings in first place, second place, and so on. When that process was complete, the wines were unbagged from last place to first place. Much discussion ensued. Only as each wine was unbagged did the recording member and all other tasters discover the maker, appellation, year, price, and other details of each wine. The tastings were double blind; no taster knew the letter label corresponding to any maker, appellation, year, or price until after the wines were ranked.

III. Randomness, Preferences, and Value

First, if a wine's position in a flight has an effect on its scores, then any conclusions about randomness, preferences, and value must take a wine's serial position into account. Does a wine's serial position have an effect on how it is scored? Is there flight-position bias? If palate fatigue, for example, causes scores on wine F to be more random, or higher or lower, than scores on wine A, then this change must be considered when comparing wines.

Each taster (t) assigns one of six ranks (r) to each wine (w) in each tasting event (e), $r_{t,w,e} \in (1, 2, 3, 4, 5, 6)$. Assigning or sampling without replacement, the mean rank ($\mu_{t,.,e}$) assigned by each taster to its six wines is 3.500, and the variance ($\sigma^2_{t,.,e}$) in those ranks is 2.917. For each wine in a tasting event, if ranks were random, the expected value of the ranks' mean ($\mu_{.,w,e}$) and variance ($\sigma^2_{.,w,e}$) are also 3.500 and 2.917. Thus, if flight-position bias exists, then the mean rank of wines in at least two positions ($\mu_{.,w,.}$ for two w) must be significantly different from 3.500.[3] Students' t-tests are employed to test the hypothesis that the observed position means $\mu_{.,w,.}$ differ from 3.500. Results, including t-test two-tailed p-values, appear in Table 1. The p-values show that none of the means differ from 3.500 with even a 75% level of significance. Accordingly, a wine's position in the flight (A, B, C, D, E, F) does not appear to affect its mean rank in a tasting.

The finding above is consistent with tasting experience. Some tasters proceed from A through F, some from F through A, some jump around, some cleanse between wines with water or bland food, some spit, and most retaste to differentiate between wines that they plan to rank closely. The potential for palate fatigue is something that experienced tasters manage to minimize flight-position bias.

---

[3] Assignment or sampling without replacement means the total of ranks must equal 21 thus position bias, if any, must be evident in at least two compensating positions. As a check on the underlying data, the sums of the means in Tables 1 and 2 are all 21 and 21/6 = 3.500.

Table 1
Mean Ranks and Flight-Position Bias

Wine	Tasting	Tasters, no.	A	B	C	D	E	F	Total
Mean of Ranks									
Trebbiano	1	9	5.00	3.56	2.44	5.44	1.89	2.67	
Calif. Sparkling Wine	2	14	3.07	3.93	3.79	2.50	4.21	3.50	
Pinot Noir, North Coast	3	19	2.47	3.47	3.58	3.47	4.63	3.37	
Nahe Rieslings	4	18	2.83	3.50	3.78	3.67	4.00	3.22	
Nero D'Avola	5	15	3.60	3.33	2.93	4.53	3.73	2.87	
Pinot Noir, Oregon	6	18	3.94	3.06	2.44	3.89	3.61	4.06	
Pinot Noir, Central Coast	7	16	3.50	3.50	5.13	3.19	2.06	3.63	
Sancerre	8	11	3.27	4.45	3.55	4.18	3.18	2.36	
Beaujolais	9	14	2.29	3.29	3.36	3.57	3.86	4.64	
Position									
Mean			3.33	3.57	3.44	3.83	3.46	3.37	
Variance			0.60	0.15	0.59	0.63	0.78	0.44	
Total wines tasted									54
Total "tastes"									804
Position Mean = 3.500									
T-test, *p*-value			0.56	0.64	0.84	0.28	0.91	0.59	

Turning to randomness and preferences, most of the authors noted in the Introduction have made estimates of the randomness in wine ratings. Their results vary, and some support the position that wine ratings, even by judges, are no better than random (Ashenfelter, 2006; Cliff and King, 1997; Hodgson, 2008; Sample, 2011; and, most recently Quandt, 2012, concerning the Judgment of Princeton). Variance in the subject data appears to support those results, however, a closer look shows that a general conclusion that wine rankings are substantially random is flawed. As shown by further tests, in the subject data and at a 95% level of significance, some aspects of tasters' rankings do not appear to be random.

Data on group mean rank $\mu_{.,w,e}$, variance in ranks $\sigma^2_{.,w,e}$, and skewness ($\gamma_{1.,w,e}$) in ranks for each tasting, arranged in order of $\mu_{.,w,e}$ from first to last place, are listed in Table 2. The means $\mu_{.,w,e}$ rise from left to right and average 3.500. The variances $\sigma^2_{.,w,e}$ appear shift-invariant. Skewnesses $\gamma_{1.,w,e}$ generally turn from positive to negative, appear shift invariant, and average nearly zero. Those results are characteristics of the moments and due to assignment or sampling without replacement.

The data in Table 2 have implications for the relative roles of preferences and randomness in tasters' rankings. A simple test suffices to show that the rankings summarized in Table 2 are not entirely random. Returning to $r_{t,w,e} \in (1, 2, 3, 4, 5, 6)$, if rankings were random, then the expected value of $\mu_{.,w,e}$ would be 3.500. Random

Table 2
Mean, Variance, and Skewness by First to Last Place

	Tasting Number	Place, determined by Mean of Ranks						Mean
		First	Second	Third	Fourth	Fifth	Sixth, Last	
Mean of Ranks								
Trebbiano	1	1.89	2.44	2.67	3.56	5.00	5.44	
Calif. Sparkling Wine	2	2.50	3.07	3.50	3.79	3.93	4.21	
Pinot Noir, North Coast	3	2.47	3.37	3.47	3.47	3.58	4.63	
Nahe Rieslings	4	2.83	3.22	3.50	3.67	3.78	4.00	
Nero D'Avola	5	2.87	2.93	3.33	3.60	3.73	4.53	
Pinot Noir, Oregon	6	2.44	3.06	3.61	3.89	3.94	4.06	
Pinot Noir, Central Coast	7	2.06	3.19	3.50	3.50	3.63	5.13	
Sancerre	8	2.36	3.18	3.27	3.55	4.18	4.45	
Beaujolais	9	2.29	3.29	3.36	3.57	3.86	4.64	
Variance of Ranks								
Trebbiano	1	0.99	1.80	1.78	1.14	0.89	0.47	
Calif. Sparkling Wine	2	2.68	1.49	2.54	3.31	2.92	2.60	
Pinot Noir, North Coast	3	1.72	3.39	2.35	2.88	2.45	2.34	
Nahe Rieslings	4	2.58	3.95	3.36	1.67	2.40	2.67	
Nero D'Avola	5	2.65	2.86	1.96	3.17	3.00	1.98	
Pinot Noir, Oregon	6	2.02	2.50	2.57	3.21	2.61	2.61	
Pinot Noir, Central Coast	7	2.43	1.53	1.38	3.13	2.11	2.11	
Sancerre	8	1.32	4.33	3.65	1.34	2.51	1.52	
Beaujolais	9	2.20	2.20	3.37	1.82	1.98	2.94	
Skewness of Ranks								
Trebbiano	1	1.09	0.64	0.42	1.16	(0.96)	(1.01)	
Calif. Sparkling Wine	2	1.04	(0.42)	0.24	(0.28)	(0.45)	(0.62)	
Pinot Noir, North Coast	3	0.82	0.28	(0.33)	(0.20)	(0.29)	(0.66)	
Nahe Rieslings	4	0.65	0.31	(0.12)	0.53	(0.77)	(0.17)	
Nero D'Avola	5	0.34	0.58	0.30	(0.18)	(0.23)	(0.64)	
Pinot Noir, Oregon	6	1.01	(0.01)	(0.18)	(0.01)	(0.33)	(0.62)	
Pinot Noir, Central Coast	7	1.52	0.04	(0.39)	0.26	(0.21)	(2.14)	
Sancerre	8	0.81	0.35	0.18	(0.95)	(0.35)	(0.39)	
Beaujolais	9	1.21	(0.12)	(0.05)	(0.27)	0.11	(0.89)	
Summary								
Mean of Means		2.41	3.08	3.36	3.62	3.96	4.57	3.50
Mean of Variances		2.07	2.67	2.55	2.41	2.32	2.14	2.36
Mean of Skewnesses		0.94	0.19	0.01	0.01	(0.39)	(0.79)	(0.01)

Table 3
T-Statistics for Mean of Ranks Deviation from 3.5

	Tasting Number	Place, determined by Mean of Ranks						
		First	Second	Third	Fourth	Fifth	Sixth, Last	Total
T-Statistics								
Trebbiano	1	4.86	2.36	1.88	0.16	4.77	8.52	
Calif. Sparkling Wine	2	2.29	1.31	0.00	0.59	0.94	1.66	
Pinot Noir, North Coast	3	3.41	0.31	0.07	0.07	0.22	3.23	
Nahe Rieslings	4	1.76	0.59	0.00	0.55	0.76	1.30	
Nero D'Avola	5	1.51	1.30	0.46	0.22	0.52	2.84	
Pinot Noir, Oregon	6	3.15	1.19	0.29	0.92	1.17	1.46	
Pinot Noir, Central Coast	7	3.69	1.01	0.00	0.00	0.34	4.48	
Sancerre	8	3.28	0.51	0.39	0.13	1.43	2.57	
Beaujolais	9	3.06	0.54	0.29	0.20	0.95	2.49	
T-Statistics, # > 2.0								
Tasting Results (above)								28%
Monte Carlo Results								8%

rankings would cause the observed $\mu_{.,w,e}$ to cluster about 3.500. The *t*-statistics in Table 3 show that 28% of the $\mu_{.,w,e}$ differ from 3.500 by over two sample-size-adjusted standard deviations.[4] These results imply strong non-random results for the highest- and lowest-ranked wines in particular.[5]

Before proceeding, the results in Table 3 were compared to the results of a random-ranking Monte Carlo simulation, and those results are summarized in the last line of Table 3.[6] After 10,000 iterations with random wine rankings, approximately 8% of the *t*-statistics were greater than 2.0.[7] Some of the apparently significant tasting results may be illusory, and this is explained after more detailed analysis.

The skewnesses $\gamma_{1.,w,e}$ in Table 2 and the significantly non-random results in Table 3 show that the distribution of ranks $r_{t,w,e}$ does not fit the moments of a random distribution. This finding is consistent with tasting experience. Many experienced tasters have some preferences in common; they agree on their relative assessments of a wine's color, clarity, finish, and general quality. While tasters may agree on some things, they are also idiosyncratic and often differ on their assessments of, and preferences concerning, attributes such as nose, varietal character,

---

[4] Absolute *t*-statistic values appear in Table 3.
[5] Even for first- and last-place, statistically insignificant results are possible if rankings are random.
[6] The author is indebted to an anonymous reviewer who suggested a Monte Carlo simulation.
[7] For each iteration of the Monte Carlo simulation, a random rank between 1 and 6 was assigned, without replacement, to each wine for each taster. 804 random wine rankings for each of 10,000 iterations.

Table 4
Paired *T*-Tests for a Difference in Means, % of wines

	Place, determined by Mean of Ranks					
	First	*Second*	*Third*	*Fourth*	*Fifth*	*Sixth, Last*
Tasting Results, *p* < 0.05						
First Place						
Second	33					
Third	33	0				
Fourth	89	11	0			
Fifth	89	11	11	11		
Sixth, Last Place	100	78	56	56	22	
Monte Carlo Results, *p* < 0.05						
First Place						
Second	2					
Third	8	0				
Fourth	19	2	0			
Fifth	38	11	2	0		
Sixth, Last Place	67	38	19	8	2	

acid, oak, char, mineral, citrus, sweetness, and balance. They have some uniform preferences but diverge on idiosyncratic preferences that are rooted in the variance in people's taste buds and neurology. For example, one taster may prefer big, oaky, buttery Chardonnay, and another may prefer the austere result of stainless steel. Although both tasters may agree that both wines are well made, they may rank the wines differently due to their idiosyncratic preferences. Neither ranking is random and without regard for any preferences at all.

If all tasters ranked a wine the same, preferences in this case would be uniform alone, and variance $\sigma^2_{.,w,e}$ would be zero. Variance can only be due to idiosyncratic preference-based rankings and randomness. The distribution of rankings $r_{t,w,e}$ is thus a mixture distribution, and its unobservable components are the degenerate distribution of uniform preferences, the distribution of idiosyncratic preferences, and the distribution of random draws. Consequently, tests applied to the variance of the mixture distribution do not isolate the effects of variance due to the random component alone.

A closer look at rankings $r_{t,w,e}$ dissects where uniform preferences, idiosyncratic preferences, and randomness may reside. Student's *t*-test was again employed to compare the mean ranks $\mu_{.,w,e}$ of the first place to second place, first place to third place, second place to third-place, and so on in each tasting. The results of these two-sample *t*-tests appear in Table 4.[8] The data show that the percentage of the tastings

---

[8] Some paired *t*-tests require variance homogeneity. An F-test is often employed to test the null hypothesis that two populations have the same variances, however, the F-test assumes that both populations are

in which each wine's mean rank $\mu_{.,w=i,e}$ differed from each other wine's mean rank $\mu_{.,w \neq i,e}$ with a one-tailed 95% level of significance.

The results in Table 4 show that the chance that a wine's mean rank $\mu_{.,w=i,e}$ is significantly different from another wine's mean rank $\mu_{.,w \neq i,e}$ increases with the difference between their mean ranks. Close calls are more likely to be separated by randomness than first from last place. For example, the mean rank of the first-place wine $\mu_{.,w=1,e}$ was different from the mean rank of the second-place wine $\mu_{.,w=2,e}$ in 33% of the tastings at a 95% level of significance.[9] The percentages increase from top to bottom and from right to left, the more distance between ranks, the more statistically significant the difference. The means of the first-place and last-place wines were different with 95% level of significance in all the tastings.

In contrast to the *t*-tests for the tasting results described above, the Monte Carlo simulation results in Table 4 show that randomness could account for a substantial portion of the apparent differences between wines.[10] Employing a wine's mean rank $\mu_{.,w,e}$ to determine the winner, loser, and other positions of wines in a flight, a common wine-tasting procedure, is prone to Type I error false positive results.[11] Randomness can differentiate any one set of ranks from another and then ordering those sets by their means can yield false differences. The more sets, the more the statistical significance of the differences between sets that are ordered farther apart. Rerunning the Monte Carlo simulation without reordering the wines changed all the simulation results in Table 4 to 14%. This residual 14% is due to the combined effects of sampling without replacement and the remaining probability that a difference in ranks can appear statistically significant.[12]

The Monte Carlo results do not mean that randomness accounts for, as an example, 67% of the difference between the first- and last-place wines. The observed ranks $r_{t,w,e}$ have a mixture distribution, and random rankings are one mixture component. Variance in observed ranks is due to both idiosyncratic preferences and randomness. The distribution of idiosyncratic preference rankings and the mixture weights on both idiosyncratic preferences and randomness are unknown. The observed variance is not decomposed; the contribution of randomness to that

---

normally distributed. Levene's Test is better suited to non-normal and skewed data (Brown and Forsythe, 1974; Levene, 1960). Median-based Levene's Test results for tasting events 1 through 9 are respectively 0.66, 0.87, 0.57, 0.22, 0.88, 0.62, 0.75, 0.08, and 0.74. Due to the unreliability of these results and the highly skewed distributions, *t*-tests allowing for variance heterogeneity are employed in Table 4.

[9] Recall that 33% means three out of nine tasting events.

[10] The Monte Carlo results in Table 4 are from the same simulations as those in Table 3, 804 random wine rankings for each of 10,000 iterations.

[11] Type I error is the failure to accept a true null hypothesis; it is a false positive indication of a difference. In this case, it is a false indication that there is a difference between wines' means.

[12] These factors were individually tested with Monte Carlo simulations. Other simulations showed that both the effects of sampling from a discrete bounded distribution and increasing the number of tasters had no significant effect on results.

Table 5
Correlation Between $/Bottle Cost and Rank

	Tasting Number	Total of Ranks	# of Indicated Ranks					
			First	Second	Third	Fourth	Fifth	Sixth, Last
Correlation Coefficient								
Trebbiano	1	(0.71)	0.82	0.60	0.00	(0.34)	(0.45)	(0.60)
Calif. Sparkling Wine	2	(0.69)	0.34	0.63	0.27	(0.36)	0.22	(0.77)
Pinot Noir, North Coast	3	(0.67)	0.68	(0.41)	0.39	0.54	0.25	(0.93)
Nahe Rieslings	4	0.07	(0.39)	0.46	0.35	0.29	(0.62)	0.41
Nero D'Avola	5							
Pinot Noir, Oregon	6	(0.14)	0.24	(0.13)	0.00	0.23	(0.69)	0.15
Pinot Noir, Central Coast	7	(0.22)	(0.19)	(0.03)	0.64	0.63	0.36	(0.69)
Sancerre	8	(0.11)	0.19	(0.30)	0.33	(0.09)	(0.39)	0.05
Beaujolais	9	0.65	(0.77)	(0.16)	0.50	(0.11)	0.53	0.17
Mean		(0.23)	0.12	0.08	0.31	0.10	(0.10)	(0.28)

variance is unknown. The Monte Carlo results in Table 4 show the substantial portion of variance that could be due to randomness, but they are not an estimate of what is due to randomness. This finding illuminates the criticism and umbrage that ferment in the popular wine press when analysts assert that wine rankings are random (recently Shaw, 2011).

Now, to value. Again, only the price range of the wines may have been disclosed during the opening remarks to the tasting. The tasting protocol was double blind, and no taster knew the price of any particular wine until the group had finished its rankings. The correlation coefficients between price per bottle and each of the total of ranks, the number of first-place ranks, the number of second-place ranks, and so on appear in Table 5.

First, the correlation between price and Total of Ranks is negative in Table 5 because more-preferred wines receive lower ranks. Consequently, the correlation between price and preference, the inverse of mean rank magnitude, has a weakly positive mean of 0.23. The correlations between price and the numbers of first place, second place, and so on ranks show high variation. Only the third-place wine's correlation to price is uniformly positive, and there is weak evidence that more last-place wines tend to be relatively expensive and first-place wines tend to be relatively less costly. In this respect, bargains seem to exist. Wine prices may be inefficient, or they may to some extent be efficient for idiosyncratic differences between tasters. Wines with certain taste profiles may be priced at a premium and worth it to idiosyncratic tasters who seek and value those profiles. Here again, results are based on correlations between price and a mixture distribution. They are not correlations between price and the nonrandom mixture component distributions.

IV. Conclusions

T-tests show that, in the subject tastings, a wine's position in a flight does not have a significant effect on the wine's mean rank. The distribution of ranks is a mixture distribution and its components are the distributions of uniform preferences, idiosyncratic preferences, and random draws. Evaluations of the variance in the mixture distribution do not isolate variance due to the random component alone. *T*-tests show that the potential for nonrandom differences between wines' ranks is positively related to the difference in wines' ranks. However, ranking procedures can lead to Type I errors that overstate the significance of differences between wines. The correlation between price per bottle and a wine's ranking is weakly positive on average but may be explained by an efficient market with idiosyncratic wine buyers.

References

Almenberg, J. and Dreber, A. (2009). *When does the price affect the taste? Results from a wine experiment*. American Association of Wine Economists AAWE Working Paper no. 35.

Amerine, M.A. and Roessler, E.B. (1976). *Wines: Their Sensory Evaluation*. San Francisco: W.H. Freeman.

Ashenfelter, O. (2006). Tales from the crypt: Bruce Kaiser tells us about the trials and tribulations of a wine judge. *Journal of Wine Economics*, 1, 173–175.

Brown, M.B. and Forsythe, A.B. (1974). Robust tests for equality of variances. *Journal of the American Statistical Association*, 69, 364–367.

Castriota, S., Curzi, D. and Delmastro, M. (2012). *Tasters' bias in wine guides' quality evaluations*. AAWE Working Paper no. 98.

Cliff, M.A. and King, M.C. (1997). Application of eggshell plots for evaluation of judges at wine competitions. *Journal of Wine Research*, 8, 75–80.

Goldstein, R., Almenberg, J., Dreber, A., Emerson, J., Herschkowitsch, A. and Katz, J. (2008). Do more expensive wines taste better? Evidence from a large sample of blind tastings. *Journal of Wine Economics*, 3(1), 1–9.

Hodgson, R.T. (2008). An examination of judge reliability at a major U.S. wine competition. *Journal of Wine Economics*, 3, 105–113.

Lecocq, S. and Visser, M. (2006). What determines wine prices: Objective vs. sensory characteristics. *Journal of Wine Economics*, 1, 42–56.

Levene, H. (1960). Robust tests for equality of variances. In: Olkin, I. (ed.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Palo Alto: Stanford University Press. 278–292.

Pecotich, A. and Ward, S. (2010). *Taste testing of wine by expert and novice consumers in the presence of variations in quality, brand and country of origin cues*. AAWE Working Paper no. 66.

Quandt, R.E. (2012). *Winetaster on 6/08/12 with 9 judges and 10 wines based on grades*. Result for the 2012 Judgment of Princeton, http://wine–economics.org/WineTastings/Judgment_of_Princeton_no_comments.html.

Sample, I. (2011). Expensive wine and cheap plonk taste the same to most people. *The Guardian*, April 13, www.guardian.co.uk/science/2011/apr/14/expensive-wine-cheap-plonk-taste.

Shaw, L. (2011). Industry attacks psychologist for flawed taste. *The Drinks Business*, April 18, www.thedrinksbusiness.com/2011/04/industry-attacks-psychologist-for-flawed-taste-test/.

Weil, R.L. (2001). Parker v. Prial: The death of the vintage chart. *Chance*, 14(4), 27–31.

Weil, R.L. (2005). Analysis of reserve and regular bottlings: Why pay for a difference only the critics claim to notice? *Chance*, 18(3), 9–15.